Measurement variability error for estimates of volume change

James A. Westfall and Paul L. Patterson

Abstract: Using quality assurance data, measurement variability distributions were developed for attributes that affect tree volume prediction. Random deviations from the measurement variability distributions were applied to 19 381 remeasured sample trees in Maine. The additional error due to measurement variation and measurement bias was estimated via a simulation study for various components of volume change. In comparison with sampling error, the error due to measurement variation was relatively small. When biases in measurements had contradictory effects on the calculation of individual tree volume, there was little additional error, however, systematic biases produced substantial error increases. The proportion of measurement variation error attributable to diameter at breast height and tree species classification was small relative to that attributable to bole (merchantable) height and percent cull attributes, which composed the preponderance of uncertainty due to measurement variation. The greatest impacts were associated with the accretion component, which was subject to measurement variation and bias at both the initial and subsequent measurements.

Résumé : Des données d'assurance de qualité ont été utilisées pour développer les distributions d'erreurs de mesure des variables servant à prédire le volume des arbres. Des valeurs d'erreur extraites aléatoirement de ces distributions ont été appliquées à 19 381 arbres faisant l'objet de mesurages répétés dans l'état du Maine. L'erreur additionnelle causée par l'effet combiné de l'erreur et du biais des mesures a été estimée en simulant les diverses composantes de la variation du volume. Par rapport à l'erreur d'échantillonnage, l'erreur de mesure est relativement faible. L'erreur additionnelle est faible lorsque le biais des mesures a des effets opposés sur le calcul du volume d'un arbre. Cependant, l'erreur augmente de façon importante lorsque le biais est systématique. La proportion de la variation de l'erreur de mesure attribuée à la classification du diamètre à hauteur de poitrine et de l'espèce d'arbre est faible relativement aux mesures de la hauteur marchande de la tige et du pourcentage de défauts qui constituent la principale source d'incertitude à cause des erreurs de mesure. Les erreurs les plus grandes sont associées à la composante d'accroissement qui est sujette à l'effet combiné de l'erreur et du biais des mesures lors du mesurage initial et des mesurages subséquents.

[Traduit par la Rédaction]

Introduction

The primary purpose of a forest inventory is to evaluate the current status of forest resources, though there is increasing emphasis on analyzing resource trends over time. Identifying trends allows the effects of current practices on attaining a desired goal (e.g., sustainability) to be evaluated. Owing to the sample-based nature of most forest inventories, determining a significant trend in resource conditions depends on the uncertainty associated with the estimated values and the desired level of confidence in the inference.

In most analyses that include sample-based estimates, the only measure of uncertainty accounted for is the sampling error that arises from lack of complete enumeration. Other sources of error have been recognized, such as measurement error, regression error, and classification error (Cunia 1965). Researchers quantifying the effects of these additional sources of error in forestry applications have generally focused on point-in-time estimation (Gertner 1990; Gertner and Köhl 1992; Gál and Bella 1995) or the propagation of error over time using a growth projection system (Gertner and Dzialowy 1984; Mowrer 1991; McRoberts et al. 1994; Gertner et al. 1995).

Most of this research has been based on a known "true" value and measured deviations from this value; thus, the term measurement error. In nearly every case, these known values are assumed so a basis for comparison can be formed. In contrast, McRoberts et al. (1994) analyzed differences among field crews measuring the same trees without making assumptions regarding deviations from a true value. As Lischke (2001) noted, differences in values that arise from the same repeated measurement protocol are indicative of uncertainty, not error. In this context, the term "measurement variability" is appropriate as it describes differences between measurements of the same attribute.

In this paper, an approach similar to that of Canavan and Hann (2004) is taken to describe the distribution of measurement variability of several tree-level attributes that affect net volume growth estimation. Measurement variability is simulated by applying measurement variability distributions to forested sample plots measured at two points in time and evaluating the resulting effects on error of estimates of total volume change. The purpose of this research is to evaluate

Received 6 November 2006. Accepted 24 April 2007. Published on the NRC Research Press Web site at cjfr.nrc.ca on 15 November 2007.

J.A. Westfall.[1] USDA Forest Service Northern Research Station, Suite 200, 11 Campus Boulevard, Newtown Square, PA 19073, USA.
P.L. Patterson. USDA Forest Service Rocky Mountain Research Station, Suite 350, 2150A Centre Avenue, Fort Collins, CO 80526, USA.

[1]Corresponding author (e-mail: jameswestfall@fs.fed.us).

Table 1. Summary of differences between two independent measurements for diameter at breast height (dbh), bole height, and percent cull from 1838 trees in Maine, New Hampshire, Pennsylvania, and Ohio.

	n	Min.	Mean	Max.	SD	Median	IQR
dbh (cm)	1838	–5.3	–0.02	8.6	0.4	0.0	0.0
Bole height (m)	1220	–7.3	0.29	9.1	1.9	0.3	2.4
Percent cull (%)	1226	–99	–0.09	99	13	0	3

Note: SD, standard deviation; IQR, interquartile range.

the effects of measurement variation on estimates of error for components of growth: ingrowth (I), accretion (A), removals (R), and mortality (M). If the distribution of measurement variation is the same at two points in time, no bias is incurred, as both samples are affected similarly (Zschokke and Lüdin 2001). Bias would be present if the distribution means differed between the two points in time. The effects of both unbiased and biased measurement variation are considered.

Data

The two data sets analyzed are from sample plots measured by the northeastern unit of the USDA Forest Service's Forest inventory and analysis program (NE-FIA) under the annual inventory system (McRoberts 2005). One data set was used to estimate the distribution of the measurement variation. The other data set was taken from the usual cyclic remeasurement of inventory sample plots. Since annual data have only been collected for a short time, the area where remeasurement data are available is much smaller than the region where data are available for estimating the distribution of the measurement variation.

The data used to develop the distributions of measurement variation originate from Maine, New Hampshire, Pennsylvania, and Ohio. In these states, blind check data were available as part of a comprehensive quality assurance (QA) program implemented by NE-FIA (USDA Forest Service 2004). A blind check is an independent remeasurement of a sample plot that occurs within 2 weeks after the normal inventory measurement. The data available provided measurement variability information for 1838 trees. These QA data are not measurement error data, as no assumptions are made regarding a "true" value for a given measurement. The differences between the two independent measurements can be attributed to measurement error, disparity in perception, inconsistent instrument calibration, mistakes reading an instrument, faulty data recording, and (or) other factors. For this study, the variables of interest are diameter at breast height (dbh), bole height (a merchantable height associated with 10.2 cm (4 in.) top diameter), percent cull (a measure of the percentage of cull below bole height; USDA Forest Service 2004), and species identification. The following techniques were used to measure the attributes: (*i*) dbh, diameter tape; (*ii*) bole height, ocular estimate, clinometer, or laser device; (*iii*) percent cull, ocular estimate; and (*iv*) species, visual inspection. The measurement variation was calculated by subtracting the value recorded by the field crew from that recorded by the QA crew. These data are summarized in Table 1.

To evaluate the effects of measurement variation, sample plots that had been remeasured as part of the regular NE-FIA inventory cycle in Maine were used. These plots were chosen because both measurements were taken on the four-point cluster plot configuration (Bechtold and Scott 2005) in a spatially distributed sampling design (Reams et al. 2005). The plots were initially measured in 1999 and remeasured in 2004. Each of the 682 plots encompasses a land area of approximately 1/6 acre (1 acre = 0.4046856 ha). To simplify estimation, only the 611 plots that were fully forested at the time of both measurements were retained. Trees with dbh ≥12.7 cm (5.0 in.) at either measurement were used for analysis (19 381 trees). Trees were assigned to growth components based on observed history. The volume of live trees with dbh <12.7 cm at initial inventory (T1) but with dbh ≥12.7 cm at remeasurement (T2) were described as I. A was determined for trees that were measured, alive, and had a dbh of at least 12.7 cm at the time of both measurements. Removal (R) volumes were trees with dbh of at least 12.7 cm at T1 but harvested prior to T2. Volume loss due to M was determined from trees with dbh of at least 12.7 cm that were alive at T1 and dead at T2. R and M volumes were based on tree characteristics at T1. Net change (N) was defined as $N = I + A - R - M$. Individual tree volumes were calculated from dbh and bole height using the equations from Scott (1981). These estimates of gross volume were converted to net volume using the observed percent cull values. The attributes of these data are given in Table 2.

Measurement variation distributions

Most analyses of this type make the assumption that measurement errors follow a normal distribution with a zero mean for unbiased errors and a nonzero mean when a bias is present (Kangas 1996, 1998; Haara 2003). Canavan and Hann (2004) outlined the problems encountered when assuming a normal distribution. They proposed a method for deriving two-stage error distributions that more closely represent the observed data. A cumulative distribution function (CDF) is constructed for the measurement variability using a point mass at zero and separately fitting functions to negative and positive portions of the empirical CDF of the measurement variation. We illustrate the process by developing the CDF for the measurement variability of dbh. The CDFs for the measurement variability of bole height and percent cull are simply stated.

Dbh

The initial step is to determine whether the distributions differ by tree size; the data were grouped into 2.54 cm dbh classes (D_i) for trees ≤38.1 cm dbh; 5.08 cm D_i for trees 38.2–50.8 cm dbh; and a single D_i for trees >50.8 cm. Each class had at least 18 observations. The data were tested to see if the distributions for the dbh classes could be modeled using normal distributions. Results of the Shapiro–Wilk test

Table 2. Summary of data from 19 381 trees in Maine at initial measurement (T1) and subsequent remeasurement (T2).

	Min.		Mean		Max.		SD	
	T1	T2	T1	T2	T1	T2	T1	T2
dbh (cm)	7.6	12.7	20.9	21.1	85.6	85.9	8.1	8.2
Bole height (m)	1.2	1.2	8.8	9.2	29.9	29.9	4.0	3.9
Percent cull (%)	0	0	8	6	99	99	17	15
Remeasurement period (years)	3	3	4.4	4.4	5	5	0.5	0.5

(Shapiro and Wilk 1965) by D_i rejected (α = 0.20) the normality hypothesis for all classes.

Using the notation of Canavan and Hann (2004), the negative-, zero-, and positive-valued measurement variations for dbh are denoted by δ_d^-, δ_d^0, and δ_d^+, respectively. The first stage is to estimate the probability density functions, $\Pr(\delta_d^-)$, $\Pr(\delta_d^0)$, and $\Pr(\delta_d^+)$, with the first step being to determine whether $\Pr(\delta_d^-)$, $\Pr(\delta_d^0)$, and $\Pr(\delta_d^+)$ are functions of dbh. Canavan and Hann (2004) found that measurement error increased with dbh. Our data indicated that measurement variation distributions for dbh were also dependent on tree size; for example, a larger proportion of smaller trees had zero measurement difference. The $\Pr(\delta_d^-)$, $\Pr(\delta_d^0)$, and $\Pr(\delta_d^+)$ were estimated using multinomial logistic regression:

[1] $$\Pr(\delta_d^-) = \frac{1}{1+\exp(\beta_0+\beta_1 D_i)+\exp(\beta_2+\beta_3 D_i)}$$

[2] $$\Pr(\delta_d^0) = \frac{\exp(\beta_0+\beta_1 D_i)}{1+\exp(\beta_0+\beta_1 D_i)+\exp(\beta_2+\beta_3 D_i)}$$

[3] $$\Pr(\delta_d^+) = \frac{\exp(\beta_2+\beta_3 D_i)}{1+\exp(\beta_0+\beta_1 D_i)+\exp(\beta_2+\beta_3 D_i)}$$

where D_i is lower left-hand endpoint of the diameter class. The estimated parameters β_i and goodness-of-fit statistics are given in Table 3. After the model was fitted, D_i was replaced with the continuous variable D, dbh.

The second stage is to model the CDFs for negative and positive measurement variations. Since dbh is reported to the nearest 0.254 cm, the CDF for the negative measurement variations is constructed by fitting a continuous function to the data in the domain $(-\infty, -0.254)$. A modeled CDF for δ_d^- is given by

[4] $$f_{\Delta_d^-}(\delta_d^-) = \begin{cases} \Pr(\Delta_d^- \le \delta_d^-) = \beta_0|\delta_d^-|^{\beta_1} & \delta_d^- \le -0.254 \\ 1 & \delta_d^- > -0.254 \end{cases}$$

where β_0 and β_1 are estimated from the data and given in Table 4. Other equations could have been used here and for subsequent analyses; the forms chosen were based on a good fit to the data ($R^2 \approx$ 0.95–0.99).

For the positive measurement variations, δ_d^+, a modeled CDF is given by

[5] $$f_{\Delta_d^+}(\delta_d^+) = \begin{cases} 0 & \delta_d^+ < 0.254 \\ \Pr(\Delta_d^+ \le \delta_d^+) = [1-\exp(\beta_0\delta_d^+)]^{\beta_1} & \delta_d^+ \ge 0.254 \end{cases}$$

where β_0 and β_1 are given in Table 4. One might expect that the CDF would also be a function of tree size. For several different model formulations, D_i were nonsignificant predictors. We surmise the large proportion of measurement variations having values of 0.254 across all D_i resulted in no statistically significant variation due to tree size.

This approach differs from that of Canavan and Hann (2004) in that they first fitted an exponential function to the positive measurement variations and then adjusted the modeled CDF by $\Pr(\delta_d^+ = 0.254)$. In our approach, a function was fitted to the positive data with $f_{\Delta_d^+}(\delta_d^+) = 0.254$ as an estimate of $\Pr(\delta_d^+ = 0.254)$. The limitation is that $f_{\Delta_d^+}(\infty)$ may not be equal to 1.

The next step is to construct a CDF for the measurement variability, δ_d, using the probabilities of the three types of measurement variability, that is, $\Pr(\delta_d^-)$, $\Pr(\delta_d^0)$, and $\Pr(\delta_d^+)$.

[6] $$f_{\Delta_d}(\delta_d) = \begin{cases} \Pr(\delta_d^-) f_{\Delta_d^-}(\delta_d) & \delta_d \le -0.254 \\ \Pr(\delta_d^-) & -0.254 < \delta_d < 0 \\ \Pr(\delta_d^-) + \Pr(\delta_d^0) & 0 \le \delta_d < 0.254 \\ \Pr(\delta_d^-) + \Pr(\delta_d^0) + \Pr(\delta_d^+) f_{\Delta_d^+}(\delta_d) & \delta_d \ge 0.254 \end{cases}$$

We remind the reader that the above distribution is a function of dbh, that is, the probability point masses are functions of D, while the CDFs are constants with respect to dbh.

There are two main steps of the process. The first step is to determine whether the distribution of measurement variation is a function of the size class variable. This is done by forming size classes and inspecting both the probabilities of negative, positive, and zero measurement variations and the CDFs for the negative and positive measurement variations by size class. Regardless of whether there is a dependency on size class, the next step is to fit a distribution to the data. At this stage, the goodness-of-fit of the normal distribution is usually tested. Owing to the usual propensity of

Table 3. Estimated parameters for multinomial regression models (eqs. 1–3) used to predict probabilities for the measurement variations of dbh (δ_d), bole height (δ_b), and percent cull (δ_c).

Parameter	β_0	SE β_0	β_1	SE β_1	β_2	SE β_2	β_3	SE β_3
δ_d	2.0862	0.1480	–0.0389	0.0055	0.9752	0.1480	–0.0361	0.0067
δ_b	–0.4722	0.3403	–0.1069	0.0247	0.4488	0.1818	–0.0493	0.0118
δ_c	2.7616	0.2208	–0.1224	0.0111	–0.7811	0.1950	0.0205	0.0074

Table 4. Estimated parameters for models describing cumulative distribution of negative (δ^-) and positive (δ^+) measurement variations for dbh, bole height, and percent cull.

				Absolute residual	
Equation	β_0	β_1	R^2	Mean	SD
4	0.0277	–2.6154	0.99	0.0089	0.0051
5	–2.9319	0.2505	0.99	0.0079	0.0058
7	–1.9211	36.9429	0.98	0.0591	0.0471
8	–0.6993	0.0824	0.98	0.0742	0.0555
9	–5.5679	0.0508	0.97	0.0708	0.0599
10	–0.2334	0.0695	0.95	0.1237	0.1101

zero measurement difference, the two-stage construction proposed in Canavan and Hann (2004) is a logical choice for modeling the distribution. The two-stage CDF depends on three attributes: (*i*) the probability distribution of the negative, zero, and positive measurement variations (this distribution may be a function of the size class variable), (*ii*) the model CDFs for the negative and positive measurement variations (these CDFs may be functions of the size class variable), and (*iii*) the degree of accuracy to which measurements are recorded; this determines the domains of the CDFs for the negative and positive measurement variations. For bole height and percent cull, the three components are described.

Bole height

The measurement variations for bole height were divided into size classes based on height. The data were assigned to 1.5 m height classes (H_i) for trees 7.6–24.4 m tall. Trees with heights <7.6 m and >24.4 m composed the remaining two classes, respectively. Each class had a minimum of 40 observations. Shapiro–Wilk (1965) tests for normality by height class showed that 9 of the 13 classes rejected the normality hypothesis (α = 0.20).

The measurement variation distributions for bole height were observed to be dependent on tree height (H). $\Pr(\delta_d^-)$, $\Pr(\delta_d^0)$, and $\Pr(\delta_d^+)$ were modeled using eqs. 1–3, with the tree height classes H_i substituted for D_i. The parameter estimates and the goodness-of-fit statistics presented in Table 3 were used.

The CDFs modeled for negative and positive measurement variations were also a function of H. The negative measurement variations, δ_b^-, were described by

$$[7] \quad \Pr(\Delta_b^- \le \delta_b^-) = \left[1 - \exp\left(\frac{\beta_0}{|\delta_b^-|}\right)\right]^{\beta_1/H} \quad -\infty \le \delta_b^- \le -0.3$$

and the positive measurement variations, δ_b^+, were described by

$$[8] \quad \Pr(\Delta_b^+ \le \delta_b^+) = [1 - \exp(\beta_0 \delta_b^+)]^{\beta_1 H} \quad 0.3 \le \delta_b^+ \le \infty$$

Estimated parameters and goodness-of-fit statistics for eqs. 7 and 8 are presented in Table 4.

Percent cull

Initial analyses, based on the diameter classes used for dbh, indicated that the distributions of measurement variation for percent cull were dependent on dbh. The probabilities for percent cull, $\Pr(\delta_c^-)$, $\Pr(\delta_c^0)$, and $\Pr(\delta_c^+)$ were modeled using eqs. 1–3, with the parameter estimates and goodness-of-fit statistics given in Table 3. The measurement variation observed for percent cull had a higher percentage of zero values than might be expected, given that it is a continuous variable. The reason for this is some trees have no cull (0%) or are entirely cull (100%). When the QA and production crews agree on either condition, percent cull values are identical. The cumulative probability for measurement variability of the negative measurement variations, δ_c^-, is described by

$$[9] \quad \Pr(\Delta_c^- \le \delta_c^-) = \left[1 - \exp\left(\frac{\beta_0}{|\delta_c^-|}\right)\right]^{D^{\beta_1}} \quad -\infty \le \delta_c^- \le -1$$

and the cumulative probability for the measurement variability of the positive measurement variations, δ_c^+, is characterized by

$$[10] \quad \Pr(\Delta_c^+ \le \delta_c^+) = [1 - \exp(\beta_0 \delta_c^+)]^{\beta_1 D} \quad 1 \le \delta_c^+ \le \infty$$

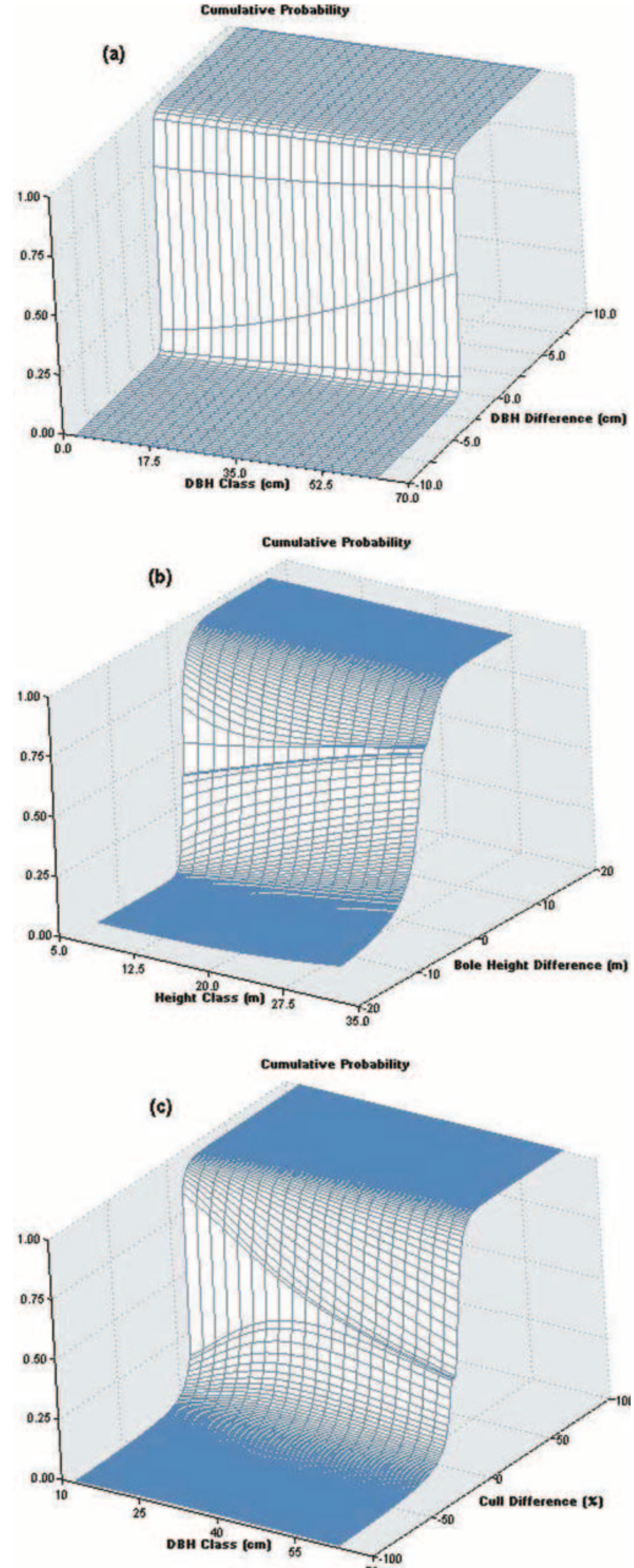
Note that eqs. 5, 8, and 10 are of identical form, as this specification worked well to describe the $\Pr(\delta^+)$. However, no form was found to work consistently well for $\Pr(\delta^-)$, indicating a single model form is inadequate for describing δ^- across all measured variables. It is assumed that measurement variation is not correlated among the attributes of interest.

The modeled two-stage CDFs for each attribute are shown in Fig. 1.

Classification variation

Classification variation occurs when items are labeled or grouped. The form of classification variation of interest here is tree species identification. NE-FIA uses 18 species groups for volume calculation, whereby each species group is associated with a specific set of volume equation coefficients (Scott 1981). If a difference in species identification results in the same coefficients, that is, different species are in the same species group, the individual-tree volume esti-

Fig. 1. Modeled measurement variation cumulative distribution function (CDF) for diameter at breast height (dbh) (*a*), bole height (*b*), and percent cull (*c*).



mate would not change. However, differences that result in classification in a different species group would produce a different estimate of volume for that tree. We considered only disagreements that produced nonequivalent volume predictions.

Table 5 shows the species identification disagreements from these data. Approximately 1.3% of the sample data showed differences in species determination that would affect volume prediction. For softwoods, the differences were attributed mostly to confusion between fir (*Abies* spp.) and spruce (*Picea* spp.). The greatest difference for hardwoods was ambiguity between maples (*Acer* spp.) and other species.

Simulated measurement variation

Simulated measurement variations were applied independently to observed sample plot data at both T1 and T2. For each tree in the Maine inventory data, measurement variation was applied for species identification, dbh, bole height, and (or) percent cull, depending on the source(s) of interest.

For dbh, bole height, and (or) percent cull, a value was chosen from the appropriate estimated measurement variation distribution and added to the observed value. This was accomplished using the standard technique of generating a random value from a uniform U(0,1) distribution and using the inverted form of the CDF to determine the estimated measurement variation. These estimated values were then rounded to the degree of accuracy of measurement, that is, 0.254 cm, 0.305 m, and 1% for dbh, bole height, and percent cull, respectively. Because extreme values of the random variable can produce untenable measurement deviations, limits of ±10.2 cm for dbh and ±12.2 m for bole height were enforced to maintain measurement variation in the approximate range of the observed data (Table 1). When this occurred, a new random value was generated until the measurement variation was within the specified limits. The limits chosen were slightly higher than those in the sample data, to account for potentially more extreme values in the parent distribution. The altered values for each attribute were also maintained within the limits prescribed for NE-FIA data collection, as field crews are permitted to enter data only within certain numerical ranges. The lower limits for dbh, bole height, and percent cull were 2.54 cm, 1.2 m, and 0%, respectively. There was no upper limit for dbh, but limits on bole height and percent cull were 30.2 m and 100%, respectively.

Variability due to tree species identification was applied randomly (via U(0,1)) using the species confusion matrix and the associated proportions of occurrence as determined from the QA data (Table 5). For example, if the QA crew species classification was red spruce and the random number was between 0 and 0.0058, the species classification was switched to balsam fir; otherwise, the species classification remained the same.

The application of measurement variation distributions to the observed data was replicated 1000 times. For each replication, the modified data were used to calculate net volume for I, A, M, and R over the remeasurement period. Changes in volume (ΔV_i) were calculated for each tree, i, on plot j (i = 1, 2, ..., n_j), (j = 1, 2, ..., n). These volumes were assigned to the appropriate change component k (k = 1 (I), k = 2

Table 5. Frequency and percent disagreement for species identification differences between the quality assurance (QA) and regular production (PC) crews that would affect prediction of individual-tree volumes.

Species identified by QA crew	Species identified by PC crew	Frequency[a]	No. of trees[b]	Percent disagreement
Balsam fir	Red spruce	2	346	0.58
Black spruce	Tamarack	1	19	5.26
Red spruce	Balsam fir	3	78	3.85
White pine	Pitch pine	2	27	7.41
Red maple	Sugar maple	8	354	2.26
Red maple	Quaking aspen	1	354	0.28
Red maple	Black cherry	1	354	0.28
Silver maple	American elm	1	4	25.00
Silver maple	Rock elm	1	4	25.00
Pin cherry	Black cherry	1	18	5.56
Black cherry	Sweet birch	1	601	1.67
White oak	Black oak	1	17	5.88
Chestnut oak	Northern red oak	1	16	6.25

[a]Number of times the species mismatch occurred.
[b]Number of trees identified as particular species by QA crew.

(A), k = 3 (R) , k = 4 (M)). Plot-level totals ($\hat{Y}_{jkh}$) for each component of change were calculated and expanded on a per-acre basis at each iteration, h, (h = 1, 2, …, 1000)

$$[11] \qquad \hat{Y}_{jkh} = \frac{\sum_i \Delta V_i}{A}$$

where A is plot size (acres). A simple random sample of the plot locations was assumed when using $\hat{Y}_{jkh}$ to estimate the population means ($\hat{\bar{Y}}_{kh}$) for each change component at each iteration

$$[12] \qquad \hat{\bar{Y}}_{kh} = \frac{\sum_j^n \hat{Y}_{jkh}}{n}$$

The generation of random numbers afforded variation among replications, thus providing an array of mean values for each component of change. The variance of these estimates represents the additional error attributable to measurement variation for each source l (l = 1 (dbh), l = 2 (bole height), l = 3 (percent cull), l = 4 (species)):

$$[13] \qquad \mathrm{Var}(\hat{\bar{Y}}_{kl}) = \frac{\sum_h^{1000} (\bar{Y}_{khl} - \hat{\bar{Y}}_{kl})^2}{999}$$

Simulations were completed for the entire data set as well as subsets of the data to evaluate the effects of sample size. Estimates of error due to sampling were obtained by calculating the variance of the estimate from the unaltered measurements.

$$[14] \qquad \mathrm{Var}(\hat{\bar{Y}}_k)_0 = \frac{\sum_j^n (\hat{Y}_{jk_0} - \hat{\bar{Y}}_{k_0})^2}{n-1}$$

Simulations were completed for application of all four sources (dbh, bole height, percent cull, and species) of additional variability simultaneously as well as for each individual source. Thus, the variance due to combined sampling and measurement variation for any desired combination can be calculated as:

$$[15] \qquad \mathrm{Var}(\hat{\bar{Y}}_k)_T = \mathrm{Var}(\hat{\bar{Y}}_k)_0 + \sum_{l=1}^{4} \mathrm{Var}(\hat{\bar{Y}}_{kl})\varphi_{kl}$$

where φ_{kl} = 1 if measurement variation from source l is included for component k and φ_{kl} = 0 if measurement variation from source l is not included for component k.

Bias

If it is assumed that the measurement variation distributions represent unbiased measurements, bias can be introduced into the simulations by shifting the central tendency from zero. Two types of bias are considered. The first is bias that occurs when the measurement biases of several variables have offsetting effects. For this analysis, a situation in which dbh, bole height, and percent cull values had the same bias trend may produce little bias in the result, that is, an increase in percent cull measurement offsets the increase in gross volume due to larger values for dbh and bole height. The other type of bias occurs when the measurement biases have a similar directional effect on net volume. For example, a positive bias for dbh and bole height combined with negative bias in percent cull would result in higher estimates of volume. The additional error due to bias is evaluated using mean squared error (MSE), which is a measure of overall accuracy:

$$[16] \qquad \mathrm{MSE}(\hat{\bar{Y}}_k)_T = \mathrm{Var}(\hat{\bar{Y}}_k)_0 + \sum_{l=1}^{4} \mathrm{Var}(\hat{\bar{Y}}_{kl})\varphi_{kl} + \left[\sum_{l=1}^{4} B(\hat{\bar{Y}}_{kl})\psi_{kl}\right]^2$$

where T is the total error (sampling and measurement varia-

Table 6. Elements of variance for estimates of change components at various sample sizes.

		Error source						
n	Change component	Sampling	dbh	Bole height	Percent cull	Species	Total	No. of trees
50	*I*	0.12002	0.00005	0.00238	0.00002	0.00000	0.12247	162
	A	4.28487	0.00961	0.59054	0.10599	0.00715	4.99816	1 183
	R	4.63963	0.00025	0.01557	0.00222	0.00024	4.65790	151
	M	1.48955	0.00028	0.01397	0.00369	0.00003	1.50751	82
	N	10.53408	0.01019	0.62246	0.11191	0.00741	11.28605	1 578
100	*I*	0.06951	0.00003	0.00121	0.00001	0.00000	0.07076	287
	A	2.02289	0.00454	0.24323	0.04735	0.00299	2.32099	2 235
	R	3.49459	0.00012	0.00583	0.00100	0.00013	3.50167	181
	M	0.91189	0.00013	0.00629	0.00290	0.00003	0.92124	184
	N	6.49887	0.00482	0.25655	0.05127	0.00315	6.81467	2 887
200	*I*	0.03223	0.00001	0.00053	0.00000	0.00000	0.03278	600
	A	1.05532	0.00191	0.12028	0.02305	0.00126	1.20182	4 784
	R	2.29736	0.00008	0.00559	0.00103	0.00003	2.30410	496
	M	0.32507	0.00006	0.00323	0.00117	0.00001	0.32953	412
	N	3.70998	0.00206	0.12962	0.02525	0.00131	3.86822	6 292
300	*I*	0.06338	0.00001	0.00042	0.00001	0.00000	0.06382	866
	A	0.71955	0.00140	0.08591	0.01390	0.00075	0.82151	6 973
	R	2.43515	0.00007	0.00409	0.00079	0.00002	2.44012	803
	M	0.31021	0.00005	0.00261	0.00103	0.00001	0.31391	656
	N	3.52828	0.00153	0.09303	0.01573	0.00078	3.63935	9 298
400	*I*	0.03925	0.00001	0.00036	0.00001	0.00000	0.03963	1 202
	A	0.56759	0.00094	0.05849	0.01080	0.00053	0.63835	9 555
	R	2.08731	0.00006	0.00298	0.00073	0.00002	2.09110	1 019
	M	0.22742	0.00004	0.00187	0.00083	0.00002	0.23017	876
	N	2.92157	0.00104	0.06370	0.01237	0.00057	2.99924	12 652
500	*I*	0.02744	0.00001	0.00024	0.00000	0.00000	0.02769	1 470
	A	0.46461	0.00082	0.04352	0.00858	0.00039	0.51792	12 002
	R	1.64319	0.00004	0.00242	0.00055	0.00001	1.64621	1 211
	M	0.17166	0.00003	0.00140	0.00057	0.00001	0.17367	1 103
	N	2.30690	0.00090	0.04758	0.00970	0.00041	2.36549	15 786
611	*I*	0.02030	0.00001	0.00020	0.00000	0.00000	0.02051	1 797
	A	0.36834	0.00066	0.03546	0.00693	0.00032	0.41170	14 839
	R	1.40284	0.00004	0.00206	0.00049	0.00001	1.40545	1 407
	M	0.14171	0.00002	0.00125	0.00046	0.00001	0.14344	1 338
	N	1.93318	0.00072	0.03897	0.00788	0.00034	1.98110	19 381

Note: *I*, ingrowth; *A*, accretion; *R*, removals; *M*, mortality; *N*, net change.

tion error), $B(\hat{\hat{Y}}_{kl})$ is the bias in the estimate of component k from source l, $\psi_{kl} = 1$ if measurement bias from source l is included for component k, and $\psi_{kl} = 0$ if measurement variation from source l is not included for component k.

Results

The results of the simulations show that sampling error composes a relatively large portion of the variance for each component of change (Table 6) when the same measurement variation distributions are applied at both T1 and T2. This is consistent with other studies on magnitude of various sources of nonsampling error (Gertner 1990; Gertner and Köhl 1992) for estimates of current values. The additional error due to measurement variation decreased proportionally with sampling error as the number of plots increased, such that the percent contribution remained essentially unchanged. Variance above sampling error alone increased by 1%–2% for *I*, 0.2%–0.4% for *R*, and about 1% for *M*. Variance of the estimate for *A* increased by 10%–14%. This result is of particular interest as *A* is the only component with measurement variation applied at both measurements (*M* and *R* only at T1 and *I* only at T2). The variance for *N* ($I + A - R - M$) increased by an average of about 4%.

Bole height was the largest measurement variation error source for each component. This result is caused primarily by the fairly wide distribution of measurement variation and the sensitivity of the volume prediction equations to changes in bole height. For all change components except *I*, percent cull was the second largest contributor to measurement variation error. Variation in percent cull values could result in large differences in the net volumes of individual trees and $\hat{Y}_{jkh}$. The low level of variation in dbh measurements resulted in little additional error (roughly 1.7%, on average, of the total measurement variation error). Variation in dbh did produce slightly more error than variation in percent cull for *I*. The contribution of additional error from variability in tree species identification also was negligible as vol-

Table 7. Elements of mean squared error (MSE) for offsetting and compounding bias situations at the T1 measurement for each component of growth.

				Offsetting bias		Compounding bias	
Change component	Estimate (m^2/ha)	Sampling error	Measurement error	Bias	MSE	Bias	MSE
I[a]	2.54	0.02	0.001	0.000	0.02	0.00	0.02
A	15.41	0.37	0.048	0.760	0.99	4.89	24.28
R	9.83	1.40	0.003	0.164	1.43	–0.37	1.54
M	6.15	0.14	0.002	0.097	0.15	–0.27	0.21
N	1.97	1.93	0.054	0.499	2.24	5.52	32.45

Note: For an explanation of abbreviations see Table 6.
[a]Not affected by bias at T1 measurement.

ume changes induced by different volume equation coefficients were relatively small compared to those induced by other sources.

As mentioned previously, the application of identical measurement variation distributions at two points in time affects both samples similarly, such that any bias cancels out. However, we wanted to evaluate the effects of bias that may occur as a result of changes in field personnel, measurement instruments, measurement protocols, etc. Comparison between simulations using all sources of measurement variation and those where only a single source was targeted indicated that the individual sources were additive. As such, the effects of bias were only evaluated for measurement variation from all four sources (dbh, bole height, percent cull, and species classification). As noted by Gertner (1990), the contribution of bias remains constant regardless of sample size, so results are based on all plots in the data. The impact of measurement bias at T1 was evaluated via the calculation of MSE as given by eq. 16. Bias occurring at T1 affects all components except *I*. Table 7 shows the components of MSE and the effect of measurement bias on overall error for both offsetting and compounding bias situations. The offsetting bias was based on bias of 1% dbh, 5% bole height, and 5% percent cull. Similarly, the compounded bias was analyzed using 1% dbh, 5% bole height, and –5% percent cull. Bias was calculated as the difference between the estimate based on the observed data and the mean estimate over 1000 replications in which bias was introduced.

Discussion

With the possible exception of *A*, the additional error due to measurement variation is relatively minor and would rarely affect the determination of significance of a trend for a particular component, that is, cause the test statistic to become too small to be significant. The *A* component exhibited a larger increase in variance than did other components, as it was subject to measurement variation at two points in time. The magnitude of measurement variability error for *A* may warrant consideration, especially when *A* is of particular interest or when estimating *N* in situations where *A* accounts for a large component of the variance. Although each component of change will likely be significantly different from zero, analysts may wish to incorporate the additional variance into the confidence interval associated with the estimate. For example, using the *A* component and a 100 plot sample would yield a 95% confidence interval width of ±0.282 using sampling error alone. Including the additional error due to measurement variation produces an interval of ±0.302, which is an increase in interval width of 7.1%. When the additional error is ignored, the attained confidence interval coverage is about 93.3%.

The difference in the relative importance of variation in bole height and percent cull measurements by change component was an interesting outcome of this study. These results reflect the common characteristics of individual trees in each change component and the lower limits of the attributes being measured. For *I*, variation in bole height adds a much larger component than does variation in percent cull. Percent cull on *I* trees tends to be small and the lower limit is zero. Thus, variation in percent cull has little opportunity to be smaller and most variation is positive. In contrast, variation in bole height for these trees has a much wider range and thus contributes more to the overall measurement variability. For the *M* component, the additional error contributed by percent cull variation is nearly twice as much as that for trees in the *A* and *R* components. Trees in the *M* component tend to have higher percent cull values, thus allowing more variation before the lower limit of zero is encountered. The relative source contributions for the *A* component represent a wide range of tree characteristics and also drive the results for net growth. Error due to measurement variability for dbh and species classification varies little by component because there is little inconsistency in measurement for these attributes.

Including measurement bias in addition to measurement variation had different effects on overall accuracy (MSE) for each change component. When biases had an offsetting effect on the prediction of individual-tree volume, MSE increased only slightly (Table 7). The additional error due to bias was larger than the error due to measurement variation for each affected component, but was still a relatively small portion (near 11%) of the total MSE. The offsetting bias results in most additional error occurring in the *A* component, which contains a large proportion of the trees from which data were collected. This offsetting type of bias is the most subtle as it may not be apparent from the estimates that measurement bias is occurring and will likely become evident only when there is a notable shift in one or more measurement variation distributions.

When measurement biases interact to compound the bias in a systematic fashion, MSE increases substantially. This is most apparent for the *A* component, in which bias results in an MSE that is roughly 59 times larger than the variance due to sampling and measurement variation alone (from Table 7, 59 = 24.28 / (0.37 + 0.04)). In this case, the *R* and *M* components contribute more to the total bias. In fact, this overestimation works in conjunction with the underestimation for *A* to increase the bias for net growth to a level higher than that exhibited by any other individual component. This produces a MSE that is over 16 times larger than would be obtained without bias. In this situation, the measurement bias may be detected as the estimates would differ greatly from expected outcomes (assuming there is some knowledge of forest resource conditions). Bias also may be detected by comparing results with data from alternative studies. For example, the Maine Forest Service (2005) compiles data on volume of wood harvested in that state each year. These and other data sources may provide a method for detecting gross biases.

Depending on the size and direction of biases for individual variables as well as interactions between the biases of several variables, there may or may not be a significant effect on MSE. Although bias may contribute little additional error for a particular estimate, inventory managers should remain concerned, as the effects of bias may appear in other estimated attributes where the offsetting phenomenon does not occur (e.g., basal area per acre). One method of identifying bias is comparing measurements of the same variables at T1 and T2. For instance, average bole height by diameter class for undisturbed plots can be calculated and compared at both T1 and T2. Any systematic differences may indicate measurement bias.

Finally, the results presented here were based on the assumption of the independence of the measured attributes. Because the data were not collected in a controlled study, assessing the validity of this assumption would be difficult. A number of potential correlations may exist, such as among measurements from individual trees, among the trees in a plot, and (or) among field measurement personnel. Further study is recommended to evaluate the amounts of dependence among these various factors.

Conclusions

The results indicate that caution is needed to ensure that forest inventory measurements are taken without bias and with minimal dispersion of measurement variation at both measurements to obtain the most accurate estimates and appropriate confidence interval coverages for change components. This can be accomplished by developing clear and concise protocols that encourage the repeatability of field measurements. Primary emphasis should be on variables that exhibit a relatively large amount of measurement variability (e.g., bole height and percent cull in this study). Additionally, these protocols should be consistent over time such that variables are measured in the same manner at both T1 and T2. The effects of changing measurement protocols should be considered carefully before implementation. Developing a method for updating values measured in the past to reflect current practices is of particular importance; otherwise, a bias may be introduced owing to the change in protocol.

Bias may develop unknowingly owing to changes in field crew personnel, measurement protocols, and (or) training methods. This is of particular importance as the results indicate that error increases substantially when systematic biases are present, particularly for the *A* component (which often comprises a large portion of the inventory). At a minimum, change component variables that have a significant influence on volume prediction should be compared between the two measurements. A related factor may be the time interval between measurements. Longer measurement intervals suggest more changes in procedures and personnel since the previous measurement.

References

Bechtold, W.A., and Scott, C.T. 2005. The forest inventory and analysis plot design. *In* The enhanced forest inventory and analysis program – national sampling design and estimation procedures. *Edited by* W.A. Bechtold and P.L. Patterson. USDA For. Serv. Gen. Tech. Rep. SRS-80. pp. 27–42.

Canavan, S.J., and Hann, D.W. 2004. The two-stage method for measurement error characterization. For. Sci. **50**: 743–756.

Cunia, T. 1965. Some theory on reliability of volume estimates in a forest inventory sample. For. Sci. **11**: 115–128.

Gál, J., and Bella, I.E. 1995. Error assessment for a provincial timber inventory. For. Chron. **71**: 627–632.

Gertner, G.Z. 1990. The sensitivity of measurement error in stand volume estimation. Can. J. For. Res. **20**: 800–804.

Gertner, G.Z., and Dzialowy, P.J. 1984. Effects of measurement errors on an individual tree-based growth projection system. Can. J. For. Res. **14**: 311–316.

Gertner, G.Z., and Köhl, M. 1992. An assessment of some nonsampling errors in a national survey using an error budget. For. Sci. **38**: 525–538.

Gertner, G.Z., Cao, X., and Zhu, H. 1995. A quality assessment of a Weibull based growth projection system. For. Ecol. Manage. **71**: 235–250. doi:10.1016/0378-1127(94)06104-Q.

Haara, A. 2003. Comparing simulation methods for modeling the errors of stand inventory data. Silva. Fenn. **37**: 477–491.

Kangas, A. 1996. On the bias and variance in tree volume predictions due to model and measurement errors. Scand. J. For. Res. **11**: 281–290.

Kangas, A. 1998. Effect of errors-in-variables on coefficients of a growth model and on prediction of growth. For. Ecol. Manage. **102**: 203–212. doi:10.1016/S0378-1127(97)00161-8.

Lischke, H. 2001. Propagation of data uncertainty through models. *In* Swiss national forest inventory: methods and models of the second assessment. *Edited by* P. Brassel and H. Lischke. WSL Swiss Fed. Res. Inst., Birmensdorf, Switzerland.

Maine Forest Service. 2005. 2004 Wood processor report. Department of Conservation, Maine Forest Service, Augusta, Me.

McRoberts, R.E. 2005. The enhanced forest inventory and analysis program. *In* The enhanced forest inventory and analysis program – national sampling design and estimation procedures. *Edited by* W.A. Bechtold and P.L. Patterson. USDA For. Serv. Gen. Tech. Rep. SRS-80. pp. 1–10.

McRoberts, R.E., Hahn, J.T., Hefty, G.J., and Van Cleve, J.R. 1994. Variation in forest inventory field measurements. Can. J. For. Res. **24**: 1766–1770.

Mowrer, H.T. 1991. Estimating components of propagated variance in growth simulation model projections. Can. J. For. Res. **21**: 379–386.

Reams, G.A., Smith, W.D., Hansen, M.H., Bechtold, W.A., Roesch, F.A., and Moisen, G.G. 2005. The forest inventory and analysis sampling frame. *In* The enhanced forest inventory and analysis program – national sampling design and estimation procedures. *Edited by* W.A. Bechtold and P.L. Patterson. USDA For. Serv. Gen. Tech. Rep. SRS-80. pp. 11–26.

Scott, C.T. 1981. Northeastern forest survey revised cubic-foot volume equations. USDA For. Serv. Res. Note NE-304.

Shapiro, S.S., and Wilk, M.B. 1965. An analysis of variance test for normality (complete samples). Biometrika, **52**: 591–611.

USDA Forest Service. 2004. Forest inventory and analysis northeast field guide. Vol. 1: field data collection procedures for phase 2 plots. Version 2.0. USDA For. Serv. Northeast. Res. Sta., Newtown Square, Pa.

Zschokke, S., and Lüdin, E. 2001. Measurement accuracy: how much is necessary? Bull. Ecol. Soc. Am. **82**: 237–243.